

CRITERII PENTRU ALCĂȚUIREA LISTEI DE INTRĂRI ÎNTR-UN DICȚIONAR DE UZ GENERAL: PROPUNERI PENTRU DEX³

MONICA VASILEANU

Institutul de Lingvistică „Iorgu Iordan – Alexandru Rosetti” / Universitatea din București
monica.vasileanu@gmail.com

Cuvinte-cheie: *dicționar de uz general, DEX, lexicografie, corpus.*

Keywords: *dictionary for general users, DEX, lexicography, corpus.*

Reeditarea *Dicționarului explicativ al limbii române* (DEX), aflată în planul Institutului de Lingvistică „Iorgu Iordan – Alexandru Rosetti”, este un prilej de a reevalua principiile care stau la baza redactării acestui dicționar. Profitând de experiența altor dicționare ale unor limbi europene, dar și de multitudinea de instrumente electronice aflate acum la dispoziția lexicografilor, noua ediție a DEX are șansa de a reaseza dicționarul atât în plan științific, printr-o definire mai clară și mai bine argumentată a principiilor lexicografice, cât și în plan social, printr-o adaptare a formei la așteptările și obișnuințele utilizatorilor.

În articolul de față propun câteva criterii mai clare, cuantificabile, pentru stabilirea listei de intrări din DEX, pe baza experienței principalelor dicționare de uz general ale limbilor engleză, franceză, italiană și spaniolă. Alcătuirea listei de intrări pornește, în general, de la definirea publicului-țintă: cine sunt utilizatorii, care este profilul lor intelectual și lingvistic și cum vor folosi ei dicționarul (Atkins, Rundell 2008: 28–30). Așadar, pentru a putea schița criteriile de incluziune, este necesară în primul rând o discuție preliminară despre utilizatorii DEX-ului și contextele în care este utilizat acest dicționar.

1. CUI ÎI ESTE DESTINAT DEX?

În scurta prefață a primei ediții, destinatarii DEX-ului sunt definiți destul de vag: „tuturor celor care, folosind în mod curent limba română, au nevoie de informații și lămuriri noi în legătură cu vocabularul românesc actual” (DEX¹). Practic, DEX poate fi folosit de vorbitori nativi și non-nativi, tineri sau bătrâni, din orice grup socioprofesional, așadar este un dicționar de uz general, nefiind destinat unui grup anume. Totuși, limitele DEX-ului sunt definite în fragmentul citat: dicționarul conține vocabularul limbii actuale. Aceeași scurtă prefață mai menționează și înregistrarea arhaismelor și a regionalismelor care apar în operele

literare clasice, sugerând principalul context de utilizare a dicționarului: în mediul școlar. Dar contextele actuale ale folosirii DEX-ului sunt mai numeroase decât se prevăzuse în 1975, iar autoritatea dicționarului este atât de mare, încât poate schimba viața unui om.

Principalul context în care folosit dicționarul rămâne mediul școlar. La școală, copiii sunt învățați să recurgă la DEX, după cum reiese din programa școlară de limba și literatura română pentru gimnaziu (Prog.). Diverse examene iau în considerare DEX-ul pentru evaluarea candidaților. Spre exemplu, la Facultatea de Drept a Universității din București, numai sensurile cuvintelor înregistrate în DEX sunt considerate corecte, utilizarea cuvintelor cu alte sensuri fiind marcată ca greșită¹. În tribunale este folosit același dicționar pentru interpretarea unor cuvinte din lege², cu consecințe importante precum încadrarea juridică a unor fapte. Utilizarea DEX-ului în jocurile de cuvinte, deși aparent fără miză importantă, a dus la crearea site-ului www.dexonline.ro, actualmente foarte accesat: prin contribuția Federației Române de Scrabble, site-ul a ajuns să fie de interes pentru utilizatori (<https://wiki.dexonline.ro/wiki/Istoric>). Și, cu siguranță, dicționarul este folosit într-o mulțime de alte contexte publice și private mai greu de documentat.

Deciziile asupra includerii sau excluderii cuvintelor derivă din caracterul de dicționar de uz general al DEX-ului, caracter dat de larga adresabilitate, precum și din contextele de utilizare, astăzi mai numeroase decât se prevăzuse în 1975.

2. CARACTERISTICILE UNUI DICȚIONAR DE UZ GENERAL – IERI ȘI AZI

Dicționarele de uz general sunt mai greu de definit tocmai pentru că pentru utilizatori acestea sunt dicționarele prototipice, orice alt tip de dicționar fiind descris printr-o sintagmă de tipul *dicționar*+ determinant. Un dicționar de uz general este obligatoriu monolingv și este publicat într-un singur volum. El este destinat publicului larg, nefiind limitat la o categorie de persoane; din același motiv, prețul și conținutul trebuie să fie accesibile. Tipul de cuvinte nu este limitat la o singură categorie, iar intrările sunt ordonate alfabetic. În fine, poate conține informații enciclopedice, dar acestea nu sunt centrale (Béjoint 2016: 7).

În cele ce urmează, principalele caracteristici ale unui dicționar de uz general vor fi enumerate pe scurt, conform unor descrieri sintetice din literatura metalexicografică (Burkhanov 1988; van Sterkenburg 1996; Swanepoel 1996; Atkins, Rundell 2008; Béjoint 2016). În mod prototipic, un dicționar de uz general

¹ A se vedea situația verbului *a adjudeca* în subiectele din 2018, în care uzul din propoziția *Și-a adjudecat premiul cel mare* este considerat greșit (<https://drept.unibuc.ro/Subiecte-de-la-concursurile-anterioare-s103-ro.htm>, accesat ultima oară în 3.02.2022).

² De exemplu, într-un proces din 2018, Înalta Curte de Casație și Justiție a apelat la DEX pentru definiția cuvântului *rană*, iar încadrarea juridică a faptelor s-a făcut ținând cont de această definiție: <http://legislatie.just.ro/Public/DetaliiDocumentAfis/199910> (accesat ultima oară în 3.02.2022).

conține vocabularul curent, uzual al unei limbi. Deși cuvintele cele mai frecvente nu sunt de obicei căutate în dicționar, caracterul reprezentativ al operei le face indispensabile. Pe lângă cuvintele aparținând limbii standard, dicționarele de uz general mai includ cuvinte comune mai puțin frecvente, expresii și locuțiuni, dar și cuvinte care nu aparțin limbii standard: unele regionalisme, termeni de argou, cuvinte familiare, unele abrevieri, termeni tehnici și științifici – ponderea acestora din urmă este în creștere. Uzul școlar face ca în dicționar să se regăsească și cuvinte mai rare, dar care apar în cele mai importante opere literare. Unele dicționare de uz general includ și nume proprii, fie la ordinea alfabetică (în ODE, de exemplu) sau într-o anexă (cum este cazul PL, *vide infra*). Sunt excluse cuvintele rare: arhaisme, regionalisme, termeni tehnici și științifici cu un grad ridicat de specializare, cuvintele de argou. Fiecare intrare conține obligatoriu definiția fiecărui sens al cuvântului, clasificarea gramaticală și informații despre flexiune, în limbile în care aceste criterii sunt relevante. În limbile în care scrierea nu este fonetică, pronunția este indicată consecvent. Definițiile sunt precedate de indicații de uz, acolo unde este nevoie: prin diverse etichete este indicată circulația cuvântului (din punct de vedere diacronic, diatopic sau diastratic). Etimologia este prezentă în majoritatea dicționarilor de uz general, însă de obicei este indicată doar la unele cuvinte și este prezentată minimal. În fine, unele dicționare conțin și imagini sau contexte ilustrative (care pot fi fabricate de lexicografi, precum în MW sau extrase din literatură, precum în PR). Editorii pot adăuga dicționarului propriu-zis diverse anexe cu caracter enciclopedic: principalele personalități ale culturii respective, sistemele de măsurătoare în cazul țărilor care nu folosesc sistemul metric etc.

Istoria acestui tip de dicționar este corelată cu schimbările sociale și tehnologice (Cowie 2009; Béjoint 2016: 8–17). Dicționarele de uz general au apărut în secolele al XVII-lea – al XVIII-lea, fiind fie versiuni prescurtate ale unor enciclopedii, fie derivate ale dicționarilor bilingve; ulterior, au fost versiuni prescurtate ale marilor dicționare istorice. În secolul al XIX-lea au apărut dicționarele-model (*Webster, Larousse*), pentru ca în secolul al XX-lea să fie redactate marile dicționare de uz general din majoritatea limbilor europene. Dezvoltarea acestui tip de operă lexicografică se află în strânsă cu formarea unei clase de mijloc, ai cărei reprezentanți aveau acces la cărți. Odată cu creșterea gradului de alfabetizare în fiecare țară, dicționarele au devenit instrumente auxiliare ale actului didactic. Astfel, păstrarea unor informații enciclopedice sau chiar a unor nume proprii se justifica prin scopul educativ al dicționarilor. Larga difuzare a impus aceste dicționare și ca instrumente de propagare sau chiar de creare a normelor literare: de exemplu, particularitățile ortografiei americane a limbii engleze au fost cristalizate prin dicționarul lui Noah Webster.

Actualmente, dicționarele de uz general se află la un punct de cotitură. După un secol în care aceste produse asigurau succesul financiar al editurilor, acum produsul tipărit este pus în umbră de produsul electronic. Toate dicționarele de uz general descrise în secțiunea următoare au și versiuni electronice, online, care în

ultima vreme au ajuns să fie consultate mai frecvent decât versiunile tipărite ale acelorași dicționare (Kosem *et al.* 2018: 12). Această preferință a tot mai multora pentru dicționarul electronic pare a anunța sfârșitul dicționarelor tipărite (Harrison 2014).

Tradiția dicționarelor de uz general din limba română nu începe cu DEX. De la finalul secolului al XIX-lea încep să apară asemenea încercări lexicografice, precum ȘDU (1896), un lexicon care a evoluat de la un dicționar cu o puternică componentă enciclopedică spre unul preponderent lingvistic: dacă în prima ediție Șăineanu a inclus și nume proprii, precum *Afganistan*, *Atlantida*, *Atlas*, *Atlantic*, ediția a VIII-a din 1930 le elimină și păstrează numai numele comune omonime, de exemplu *atlas* „colecție de hărți”. În prima jumătate a secolului al XX-lea, cea mai reușită întreprindere de acest tip este CADE (1931). În a doua jumătate a secolului trecut, Academia Română își asumă dicționarele de uz general. Apar astfel DLRLC, coordonat de Dimitrie Macrea și Emil Petrovici, în 4 volume (1955–1958) și o versiune mai scurtă, mai accesibilă, DLRM, într-un singur volum, coordonat de Dimitrie Macrea (1958). Prima ediție a DEX a fost publicată în 1975. Alternative la DEX au apărut târziu, după anul 2000. NDU (2006) continuă dicționarul lui Lazăr Șăineanu. În aceeași perioadă este publicat și DEXI (2007). Ambele au umplut un gol – DEX-ul nu mai fusese îmbogățit, iar limba română se schimba rapid –, dar nu au dobândit autoritatea dicționarului elaborat de Academia Română.

Din păcate, principiile de alcătuire a dicționarului nu sunt menționate în prefață și nici edițiile următoare nu au beneficiat de vreun paratekst lămuritor. De fapt, din 1975, DEX nu a mai trecut printr-o revizie fundamentală. A doua ediție (1996) nu a făcut decât să adauge cuvintele din *Suplimentul* apărut în 1988 și a eliminat reminiscențele comuniste evidente: au dispărut din definiții cuvinte precum *tovarăș* și *pionier* și au fost rescrise definițiile puternic marcate ideologic, de exemplu, la *capitalism*. Tirajele ulterioare au adăugat cuvinte și sensuri (2009, 2016) și au modificat punctual câteva definiții contestate de diverse grupuri sociale (de exemplu, *jidan* în 2012). Însă ansamblul dicționarului a rămas același. Totuși, vocabularul limbii actuale din 1975 nu este același cu vocabularul limbii actuale din 2021. De aceea este necesară o revizuire profundă a listei de cuvinte, precum și a definițiilor.

3. CRITERII DE INCLUDERE A CUVINTELOR ÎN DICȚIONARE DE UZ GENERAL: CÂTEVA EXEMPLE DIN LEXICOGRAFIA INTERNAȚIONALĂ

Este destul de dificil de identificat un dicționar care să corespundă DEX-ului în lexicografia britanică, franceză sau italiană. Acest lucru se întâmplă în primul rând din cauza diferențelor instituționale. DEX este elaborat de o autoritate națională, Institutul de Lingvistică al Academiei Române, pe când în alte țări nu există o autoritate similară – unică, centrală. Singura situație comparabilă cu a DEX-ului o constituie cea din Spania, unde DLE este elaborat de 23 de academii din întreaga lume hispanofonă, ceea ce îi conferă autoritate. În rest, elaborarea

dicționarelor de uz general este descentralizată și mai degrabă efectuată în mediul editorial privat. În Franța există o Academie cu mare autoritate care impune norma literară, dar nu sub egida ei sunt elaborate principalele dicționare de uz general: acestea se realizează sub patronajul unor edituri recunoscute, fie universitare (precum PUF), fie comerciale (Éditions Larousse și Robert). În Regatul Unit al Marii Britanii și Irlandei de Nord, tot marile edituri sunt cele care publică dicționarele, uneori în colaborare cu un centru universitar. La fel arată și peisajul editorial din Italia, cu o concurență între diversele edituri care publică dicționare. În toate aceste cazuri, nu există un singur dicționar care să fie „Dicționarul”, ci o multitudine de produse lexicografice, aflate pe o piață concurențială, ceea ce stimulează inovația.

În privința alcătuirii listei de cuvinte, dicționarele consultate explică, în general, în prefață ce tipuri de intrări conțin, fără a specifica în detaliu cum au ajuns la acea listă de cuvinte. În general, istoria fiecărui dicționar oferă indicii cu privire la selecția intrărilor. Spre exemplu, PL este versiunea prescurtată a unui dicționar enciclopedic, *Le Grand Larousse*, iar actualul De Mauro este versiunea prescurtată a *Grande Dizionario della Lingua Italiana* a lui Tullio De Mauro. Dar și acestea au fost de-a lungul timpului supuse unor revizii, pentru a reflecta schimbările de vocabular; prefețele edițiilor mai noi explică, mai succint sau mai pe larg, și criteriile de includere a cuvintelor noi în dicționar.

3.1. EXEMPLE DIN LEXICOGRAFIA BRITANICĂ

Pentru limba engleză am selectat două dicționare de uz general din spațiul britanic. Selecția a fost dificilă, întrucât limba engleză beneficiază de foarte multe și variate dicționare, iar o parte dintre cele mai folosite dicționare nu se încadrează în tipologia unui dicționar de uz general. Spre exemplu, dicționarul electronic MW³, cel mai cunoscut dicționar american, se bazează pe o versiune de uz școlar, *Merriam Webster's Collegiate Dictionary*. Aceeași editură a publicat numeroase dicționare, toate fiind dezvoltări ale lucrării lui Noah Webster, dar niciuna nu corespunde exact caracteristicilor unui dicționar de uz general. Editura britanică Macmillan are în portofoliu multe dicționare, foarte consultate, dar acestea sunt lucrări destinate învățării limbii engleze (*learners' dictionary*). Am selectat pentru analiză două dicționare din lexicografia britanică, ODE și CED, care își propun să cuprindă într-un singur volum vocabularul reprezentativ al limbii engleze actuale.

ODE a fost publicat de Oxford University Press în 2010, atât în formă tipărită, cât și online. Este un dicționar într-un singur volum, reprezentativ pentru limba engleză actuală. Ediția din 2010 cuprindea aproximativ 100.000 de intrări, rezultând c. 357.000 de cuvinte, expresii și sensuri. În 2015, versiunea online a fost

³ Detaliile despre istoricul fiecăruia dintre dicționarele descrise sunt preluate de pe site-urile dicționarelor respective. Am indicat în bibliografie adresele site-urilor.

îmbogățită cu c. 80.000 de cuvinte extrase din *Oxford English Dictionary*. Pe lângă cuvintele comune, conține și o serie de nume proprii, considerate de editori parte din limbă. Ocazional, dicționarul oferă și contexte ilustrative reale, nu fabricate, dar fără trimiteri bibliografice. Alcătuirea listei de intrări s-a bazat pe cercetarea de corpus: au fost folosite *British National Corpus* și *Oxford English Corpus*, acesta din urmă însumând peste 2 miliarde de cuvinte. O altă sursă importantă o reprezintă *Oxford Reading Programme*, un grup de aproximativ 50 de cititori angajați special să semnaleze cuvinte din diverse materiale, inclusiv din cele audio-video (emisiuni de televiziune, versurile melodiilor etc.). De asemenea, editura a apelat la consultanți din toate colțurile lumii unde se vorbește engleză. Așadar, ODE se bazează pe un set de date lingvistice complex, primordială fiind analiza de corpus. Versiunea online este îmbogățită de mai multe ori pe an: în 2021 îmbogățirea s-a petrecut la fiecare 6 luni, fiind adăugate mai mult de 400 de cuvinte, expresii și sensuri noi de fiecare dată, după cum raportează editura pe site-ul propriu (<https://languages.oup.com/about-us/annual-updates/>, accesat în 20.11.2021).

Publicat pentru prima dată în 1979, CED a fost primul dicționar britanic într-un singur volum care cuprindea și toponime și antroponime, considerate parte din limba curentă la fel ca numele comune. Ediția a 13-a, pe care am consultat-o, a fost tipărită în același an cu ODE, anume în 2010. În 2018 a fost reeditat cu adăugiri, ajungând la 725.000 de cuvinte, sensuri și expresii. Selecția intrărilor se face pe baza *Collins Corpus*, care actualmente a ajuns la 4,5 miliarde de cuvinte. Cu ajutorul acestui corpus atât de cuprinzător, în continuă creștere, sunt detectate cuvintele noi, compusele și expresiile noi, sensurile sau utilizările noi. În plus, utilizatorii site-ului dicționarului pot sugera cuvinte pe care și-ar dori să le găsească în dicționar.

Așadar alcătuirea listei de cuvinte din cele două dicționare britanice alese se bazează pe analiza de corpus. Prefetele dicționarelor citate nu conțin explicații mai detaliate, care să indice și o cuantificare a datelor din corpus. Cele două lucrări discutate mai sus prezintă diferențe între forma tipărită și versiunea electronică: cuvintele noi sunt introduse pe site, iar versiunile tipărite se succedă la intervale de timp destul de mari (mai mulți ani). Așadar editorii au decis să decupleze dicționarul imprimat de cel electronic.

3.2. EXEMPLE DIN LEXICOGRAFIA FRANCEZĂ

Cele mai populare dicționare de uz general ale limbii franceze reprezintă variantele prescurtate ale unor dicționare mult mai cuprinzătoare. PR a apărut prima dată în 1967, fiind prezentat publicului ca o variantă mai concisă a *Dictionnaire alphabétique et analogique de la langue française*, de Paul Robert, publicat în 6 volume. Popularitatea dicționarului a crescut odată cu preluarea conducerii de către Alain Rey. Începând din 1996, PR a avut și o versiune electronică, pe CD, iar din 2009 poate fi consultat și prin site-ul propriu, pe baza

unui abonament. Pentru fiecare intrare se definesc sensurile, se prezintă pronunția și principalele informații gramaticale; în plus, dicționarul prezintă sinonimele și antonimele cuvântului-titlu. PR nu conține nume proprii, acestora fiindu-le dedicat un alt dicționar. În prezent, PR este revizuit și republicat anual. Spre exemplu, ediția 2022, apărută în 2021, conține 300.000 de cuvinte și sensuri, 35.000 de citate, 75.000 de indicații etimologice. Prezentarea acestei ediții nu precizează câte cuvinte noi au fost introduse (față de anul precedent), ci doar ilustrează principalele situații. Evenimentul principal care a determinat o schimbare majoră a vocabularului a fost pandemia de SARS-CoV-2, dar și alte domenii, precum sportul, showbizul și gastronomia au impus noi cuvinte. Între acestea, se regăsesc și o seamă de anglicisme. Se mai menționează și introducerea unor cuvinte specifice doar anumitor regiuni francofone din afara Franței. Prezentarea indică și faptul că aceste cuvinte noi au fost incluse în versiunea online încă din 2020.

Aproape în oglindă se prezintă PL, cu deosebirea că acesta derivă, prin avataruri intermediare, dintr-un dicționar enciclopedic, fapt care se reflectă în structura lui. Dicționarul propriu-zis este urmat de câteva „pagini roz”, care cuprind expresii și locuțiuni, majoritatea în latină, și de o secțiune cu nume proprii, așadar cu caracter enciclopedic. La fel ca PR, PL este republicat cu modificări anuale. Ediția pentru anul 2022, apărută în 2021, conține aproximativ 63.500 de intrări, cu 125.000 de sensuri, plus 20.000 de expresii și 28.000 de nume proprii. Odată cu volumul tipărit, cumpărătorii primesc și o cheie de acces la forma electronică, mai cuprinzătoare, cu 80.000 de intrări. Prefața ediției menționează introducerea a 150 de cuvinte în dicționar, majoritatea legate de criza sanitară. Introducerea acestor cuvinte nu înseamnă neapărat că ele sunt noi: în cazul unor lexeme precum *asymptotique* s-a petrecut o modificare a circulației, uzul curent calificându-le pentru includerea într-un dicționar de uz general. De asemenea, editorii au introdus expresii și locuțiuni noi, precum și afixe noi (de exemplu, *corona-*).

Din păcate, nu am găsit explicații detaliate ale modalității de culegere, de identificare și de evaluare a cuvintelor care au îmbogățit aceste ediții anuale. Dar exemplele indică o actualizare rapidă a dicționarilor: editorii nu au așteptat să vadă care dintre aceste cuvinte sunt efemeride, ci le-au introdus în ediția anuală a dicționarului. Cu siguranță, acesta este și punctul forte al edițiilor anuale. Nu am găsit nicio mențiune referitoare la eliminarea cuvintelor din aceste ediții anuale, dar, dat fiind că dimensiunea dicționarilor rămâne constantă, acest lucru trebuie să se petreacă la un moment dat.

3.3. EXEMPLE DIN LEXICOGRAFIA ITALIANĂ

Și din lexicografia italiană selecția a fost dificilă: dintre multele dicționare monolingve într-un singur volum, lucrări excelente, cu tradiție, m-am oprit cu greu la trei lucrări.

De Mauro reprezintă versiunea prescurtată a *Grande dizionario italiano dell'uso*, redactat de Tullio De Mauro în 6 volume. Tipărit în 2000, dicționarul într-un

volum cuprinde 130.000 de intrări și 30.000 de compuse sau expresii. Prefața acestui dicționar de uz general menționează care sunt tipurile de cuvinte incluse: cuvinte fundamentale și uzuale, cuvinte „cu înaltă disponibilitate” (chiar dacă nu sunt frecvent atestate, sunt legate de realități cotidiene și pot fi folosite oricând), termeni tehnici și științifici, cuvinte regionale, cuvinte rare, dar prezente în literatura din secolul al XX-lea, cuvinte ieșite din uz, dar prezente în alte dicționare, prefixe, sufixe și elemente de compunere, cuvinte neadaptate, abrevieri și sigle.

Unul dintre cele mai populare dicționare este Zingarelli, publicat de editura Zanichelli și bazat pe *Vocabolario della lingua italiana*, de Nicola Zingarelli, apărut prima dată în 1917. Începând cu 1994, este revizuit și îmbogățit anual pentru a ține pasul cu schimbările lingvistice. Ediția 2022, de exemplu, conține aproximativ 145.000 de intrări cu 380.000, dintre care 1000 sunt cuvinte, sensuri și expresii nou introduse. Coordonat actualmente de Mario Cannella, dicționarul conține, pe lângă lexicul standard, și termeni tehnici care au intrat în limba comună, termeni politici și administrativi care se difuzează rapid, precum și neologisme de diferite feluri. Este, de altfel, cel mai deschis dicționar italian de uz general în privința neologismelor (Marello 2020). Pe lângă îmbogățirea permanentă și rapidă, dicționarul este în permanență revizuit. Revizuirea se face pe domenii (două-trei domenii pe an). Pentru a păstra o dimensiune confortabilă a dicționarului, editorii elimină cuvintele rare, inuzitate dacă nu au valoare culturală (CB).

O strategie diferită a fost adoptată de editorii de la Treccani. Pe baza lucrării *Vocabolario della lingua italiana*, publicat între 1986 și 1994 sub coordonarea lui Aldo Duro, a fost redactat un dicționar de uz general, într-un singur volum, *Il Nuovo Treccani*, apărut în 2018. Editorii și-au propus să ofere publicului un dicționar simplu, cu metalimbaj redus la minimum. De asemenea, au eliminat cuvintele neclare, hapaxuri păstrate din inerție în dicționare. Această lucrare însă face parte dintr-un pachet de 4 volume, care cuprinde și o gramatică, un dicționar de sinonime și antonime și un dicționar de neologisme: *Il Nuovo Treccani, Neologismi. Parole nuove dai giornali* (2008–2018) cuprinde aproximativ 3500 de cuvinte, bazat pe o bază de date cu neologisme extrase din ziarele centrale și locale. Așadar, Institutul Treccani a optat pentru o separare a cuvintelor recente într-un volum dedicat.

Situația celor trei dicționare menționate este aproximativ aceeași: ele se bazează pe dicționare mai ample și mai vechi. Alcătuirea listei de intrări se va fi făcut, în dicționarele-bază, prin metode tradiționale, în principiu prin fișarea materialului consultat. Dar, cu trecerea timpului, noile tehnologii au adus posibilitatea reviziei și a îmbogățirii listei de cuvinte. Precizări mai clare am găsit pentru Zingarelli: coordonatorii specifică faptul că includerea sau excluderea unui cuvânt din dicționar se bazează pe uz, fără a oferi însă detalii cuantificabile (CB).

3.4. UN EXEMPLU DIN LEXICOGRAFIA SPANIOLĂ

Singurul dicționar dintre cele consultate care se află într-o situație oarecum similară cu DEX este *Diccionario de la lengua española*, care este redactat de o autoritate lingvistică, Real Academia Española, acum în colaborare cu Asociación de Academias de la Lengua Española. Prima ediție a apărut în 1780 și reprezenta versiunea prescurtată – în special prin eliminarea citatelor ilustrative – a *Diccionario de autoridades*, primul dicționar academic spaniol. Cea mai recentă ediție a fost publicată în 2014 și conține 93.111 intrări cu 195.439 de sensuri. DLE cuprinde lexicul uzual al limbii spaniole din Spania și America Latină, Filipine și Guinea Ecuatorială: cuvinte uzuale (din secolul al XVI-lea până în prezent pentru a ajuta cititorii de literatură), colocviale, termeni tehnici și științifici, diminutivele și augmentativele, doar dacă au vreo particularitate semantică sau morfologică. Prefața acestei a 23-a ediții menționează ca principal criteriu de includere uzul stabil al unui cuvânt, stabilit pe baza unui corpus. DLE nu acceptă noutățile lexicale decât după ce au trecut proba timpului și se ferește să accepte efemeride lexicale.

3.5. CRITERII CUANTIFICABILE

După cum am putut constata din cele de mai sus, dicționarele de uz general sus-menționate din Franța, Italia și Spania sunt versiuni prescurtate ale unor dicționare mai mari. Alcătuirea listei de cuvinte a constat în selectarea cuvintelor dintr-o listă mai mare de cuvinte, preexistentă, pe diferite criterii, pentru a respecta limitarea de spațiu, anume un volum, și astfel a face dicționarul accesibil. Ulterior, lista de cuvinte a fost îmbogățită pentru a reflecta schimbările lexicale, în special în versiunile electronice ale dicționarilor. Cu totul diferită este situația ODE și CED, dicționarele britanice fiind lucrări autonome. Lista de intrări a fost alcătuită, în aceste două cazuri, pe baza unor corpusuri extinse. Despre corpus se vorbește ocazional și în cazul celorlalte dicționare, însă în aceste cazuri el reprezintă doar instrumentul de verificare a circulației reale a cuvintelor, nu punctul de plecare. Însă în niciunul dintre materiale nu am găsit cuantificatori preciși.

Deși nu este un dicționar de uz general, experiența *Oxford English Dictionary* poate fi utilă. Având ambiția de a cuprinde cât mai mult din lexicul limbii engleze, editorii nu exclud cuvinte. Dar măcar în privința includerii, criteriile sunt clare. Cuvintele candidate la includere sunt propuse de Oxford Reading Programme. Circulația acestor cuvinte este evaluată prin interogarea *British National Corpus*, a *Oxford English Corpus* (care conține texte datând din anul 2000 până în prezent) și a altor corpusuri sau colecții de periodice. În 2006, filtrul de includere a cuvintelor era „regula celor 5”: un cuvânt putea fi inclus în dicționar dacă apărea de cel puțin 5 ori, în 5 surse diferite, într-o perioadă de 5 ani (Maxwell 2006). Cuantificatorii s-au schimbat odată cu dezvoltarea accelerată a internetului (Diamond 2016):

a. numărul de atestări: pentru a fi inclus, un cuvânt trebuie să aibă minimum 100 de utilizări independente;

b. vârsta: pentru fiecare cuvânt trebuie să fie documentată o atestare continuă de minimum 10 ani;

c. „naturalizarea”: cuvântul trebuie să fie atestat fără glose sau explicații, dovedind că este înțeles (sau cel puțin așa se așteaptă locutorul);

d. semnificația lexicografică: pot fi incluse în dicționar și cuvinte slab atestate care au dezvoltat o familie lexicală mai bine atestată.

Bineînțeles, pentru un dicționar de uz general, criteriile sunt ușor diferite și depind de dimensiunile dicționarului și de publicul-țintă. Cu cât dicționarul este mai mic, cu atât criteriile vor deveni mai restrictive, de exemplu va fi nevoie de un număr mai mare de atestări pentru a include un cuvânt și, bineînțeles, vor fi eliminate cuvintele ieșite din uz (Maxwell 2006).

Principala metodă de identificare a cuvintelor noi o reprezintă în continuare extragerea din diverse materiale: cărți, reviste, ziare, publicații electronice. În unele redacții, această căutare este efectuată de oameni (MW, precum și Oxford Reading Programme). Însă munca a fost deja automatizată de unii editori: există deja instrumente electronice prin care se compară materialul lexical din texte cu baza de cuvinte deja înregistrate în dicționar, astfel încât cuvintele noi să fie evidențiate (Maxwell 2006).

Însă niciun criteriu nu este infailibil *per se*. În anumite situații, lexicografii pot decide să încalce limitele cantitative propuse și să includă în dicționar un cuvânt mai tânăr, dar cu largă difuzare și cu semnificație socială. Criteriile sunt instrumente ale lexicografului, nu stăpâni ai lui, iar un lingvist cu experiență poate prezice în ce cazuri un cuvânt nou are potențialul de a se difuza în limbă și îl poate include din timp în dicționar. Un exemplu în acest sens este eng. *AIDS*, care a fost înregistrat în dicționare la câteva luni după apariție (MW). În plus, un dicționar al limbii curente ar trebui să fie revizuit frecvent, astfel încât să reflecte starea curentă a vocabularului; astfel, cuvintele pot fi înregistrate mai devreme, urmând a fi eliminate la următoarea revizie dacă nu vor mai avea un uz stabil (Maxwell 2006).

4. PROPUNERI PENTRU DEX³

Fiind un dicționar de uz general, DEX ar trebui să conțină în primul rând cuvintele aparținând fondului principal lexical, precum și acele cuvinte „cu înaltă disponibilitate”, care se referă la realitățile înconjurătoare, fără a fi neapărat foarte frecvente, și cuvintele aparținând fondului cultural din diverse domenii (fie ele științe umaniste sau exacte). În plus, contextul principal de folosire, anume cel școlar, impune păstrarea unor cuvinte populare, regionale sau arhaice, însă numai în măsura în care ele se întâlnesc în operele școlare sau au dezvoltat derivate sau compuse cu mai largă circulație. În plus, pentru a reflecta stadiul curent al limbii române, DEX³ ar trebui să includă și anumite neologisme recente, cu largă

circulație, precum și termeni tehnici și științifici care au pătruns în limba curentă, mai ales în actualul context al vulgarizării științelor. De asemenea, ar trebui să cuprindă și cuvinte familiare provenite din diverse stiluri, inclusiv argotice, în măsura în care circulația lor în limba curentă poate fi dovedită. În fine, întrucât proliferarea compunerii savante nu poate fi surprinsă în DEX, principalele elemente de compunere și afixele nou detașate ar trebui definite, permițând utilizatorului să își explice formațiunile mai noi neînregistrate în dicționar (ex. *-zaur*, *-crație*). În fine, circulația în limba curentă ar putea impune și includerea unor sigle și derivate ale lor, precum și unele denumiri comerciale, marcate ca atare, în special cele care prin antonomază au devenit și cuvinte comune (ex. *adidas*).

Pentru a-și păstra caracterul accesibil, DEX³ nu ar trebui să includă cuvintele cu circulație redusă: regionalisme sau arhaisme care nu sunt atestate în literatura școlară, terminologia de strictă specialitate, împrumuturile recente care încă nu și-au dovedit viabilitatea, creațiile ocazionale strict dependente de un anumit context politic sau social, elementele de argou care nu au intrat în limba comună, majoritatea siglelor și a cuvintelor trunchiate, care sunt în general efemere, precum și numele proprii, specifice dicționarilor enciclopedice (deși unele derivate de la nume proprii, de ex. *eminescian*, ar trebui incluse).

După cum am precizat mai sus, DEX nu a trecut până acum printr-o revizie profundă. În mare, lista de intrări este aceeași din 1975: adăugiri s-au mai efectuat (fără a se explicita criteriile), dar excluderi nu. Însă din 1975 și până acum vocabularul s-a schimbat considerabil. Pentru a păstra caracterul accesibil (în primul rând sub aspectul dimensiunii) și, mai ales, reprezentativ al DEX, noua ediție trebuie să reevalueze complet lista de intrări, verificând compatibilitatea fiecărui cuvânt-intrare cu criteriile pe care echipa de redactori le va asuma.

Întrucât ne aflăm într-o perioadă de reflectare asupra principiilor care vor sta la baza noii ediții a DEX-ului, în cele ce urmează vom schița o propunere personală de asumare a unei proceduri de lucru pentru alcătuirea listei de cuvinte. Subliniem faptul că aceste idei nu au fost încă discutate în colectivul de redacție. Propunerile au la bază experiența dicționarilor de uz general descrise în secțiunea anterioară, dar țin cont și de resursele disponibile în momentul de față la Institutul de Lingvistică „Iorgu Iordan – Alexandru Rosetti”.

Așadar, pentru alcătuirea listei de cuvinte a noului DEX este esențială verificarea uzului cuvintelor (a) în limba curentă și (b) în context școlar. Considerăm necesar ca această verificare să se petreacă atât pentru cuvintele deja înregistrate, cât și pentru cuvintele-candidate pentru includere.

Verificarea circulației cuvintelor în context școlar pare la prima vedere o întreprindere simplă. S-ar putea face alcătuind un corpus cu textele școlare, care să cuprindă manualele de la toate disciplinele predate în limba română, din clasele I–XII/XIII, precum și textele integrale ale operelor literare din care sunt extrase fragmentele prezente în manualele de limba și literatura română. Însă un asemenea corpus exhaustiv este, din păcate, nerealist. Când au fost redactate primele două

ediții ale DEX-ului, programa școlară pentru limba și literatura română conținea o listă exhaustivă de lecturi, iar manualele erau unice. Acum însă programa școlară sugerează niște autori și niște texte pentru a fi parcurse (Prog.), dar lista este orientativă, nu exhaustivă. Am observat, de exemplu, că manualul de limba română din clasa a VI-a de la editura Art include și fragmente traduse din literatura universală. Practic, corpusul s-ar lărgi și cu texte traduse, ceea ce este problematic. În plus, ne aflăm în plin proces de refacere a planurilor-cadru și a programelor de studiu pentru liceu, iar manualele actuale se vor schimba și ele. Această dinamică a peisajului școlar nu permite realizarea a unui corpus exhaustiv de texte prezente în context școlar.

Mai realistă ar fi întreprinderea de realizare a unui corpus reprezentativ de texte școlare, care să cuprindă doar textele tuturor manualelor aprobate de Ministerul Educației pentru disciplinele cu predare în limba română (inclusiv fizică, chimie etc., întrucât sunt necesare pentru evaluarea circulației termenilor tehnici și științifici). În cazul manualelor de limba și literatura română, riscul este ca elevilor să li se ceară să citească integral o operă din care manualul cuprinde doar un fragment, iar elevul să nu găsească în DEX un cuvânt întâlnit în restul operei. Această problemă nu poate fi rezolvată doar între paginile DEX. O soluție la acest inconvenient ar putea fi reprezentată de decuplarea site-ului de forma tipărită a dicționarului, idee pe care o vom dezvolta mai jos.

Mai complicată este verificarea circulației cuvintelor în limba curentă. Dicționarele prezentate în secțiunea 3 au răspuns, în majoritate, la această problemă cu un răspuns unitar: analiza de corpus. Însă pentru ca analiza de corpus să fie relevantă, trebuie luat în considerare un corpus reprezentativ, ceea ce este în sine o provocare, întrucât niciun corpus scris nu poate cuprinde toată varietatea limbii vorbite, pe de o parte, iar, pe de altă parte, anumite cuvinte vor fi favorizate de evenimente sociale sau politice sau pur și simplu de disponibilitatea textelor-sursă, în vreme ce altele vor fi ocluate din aceleași cauze (Kupietz 2016; Kilgariff 2013). Pentru limba română, niciunul dintre corpusurile disponibile nu este relevant pentru DEX. Corpusul larg de texte alcătuit în cadrul Departamentului de lexicografie al Institutului de Lingvistică „Iorgu Iordan – Alexandru Rosetti” este util pentru o cercetare istorică, fiind un instrument indispensabil în redactarea *Dicționarului etimologic al limbii române* și a *Dicționarului limbii române*, dar cuprinde prea puține texte contemporane. Corpusurile orale *Interacțiunea verbală în limba română* sau *Corpus de română vorbită* sunt, prin natura transcrierii, neinterogabile, fiind foarte utile pentru recuperarea informației paralingvistice (intonație, succesiunea replicilor etc.). Cel mai mare corpus adnotat al limbii române, CoRoLa, nu este nici el reprezentativ, întrucât 90% din corpus este alcătuit din texte juridice și științifice, proporție care, cu siguranță, nu reflectă limba curentă. Corpusul oral care îl însoțește este prea mic pentru a compensa acest neajuns, iar metadatele sunt adeseori insuficiente: anumite texte sunt foarte greu de datat, trebuind efectuate căutări suplimentare. În plus, CoRoLa este un corpus finit,

publicat în 2017. În anul 2027, când este programată publicarea DEX³, deja el nu va mai fi reprezentativ pentru limba curentă din anul respectiv.

Prin urmare, este absolut necesar un corpus de presă scrisă și *new media* din ultimii 10 ani, care să reflecte mai fidel starea actuală a limbii române. Acest corpus ar trebui să includă articole din periodice, dar și postări din *social media*, de pe bloguri, discuții de pe forumuri, comentarii la articolele de presă, precum și transcrieri ale unor texte orale disponibile (de exemplu, stenograme ale ședințelor Parlamentului, transcrieri ale unor emisiuni). Un asemenea corpus ar putea fi compilat relativ repede cu ajutorul unor parteneri din presă și din domeniul editorial și din diverse instituții ale statului, precum și cu ajutorul unor instrumente electronice noi, precum API-urile unor rețele de socializare⁴. Credem cu tărie că acest corpus ar trebui asumat de Institutul de Lingvistică „Iorgu Iordan – Alexandru Rosetti” ca proiect fundamental, întrucât ar putea fi folosit pentru cercetări în toate domeniile limbii. Corpusul ar urma să fie îmbogățit în fiecare an, cu un număr de texte proporțional cu resursele (materiale, tehnologice, dar mai ales umane) disponibile în institut pentru acest proiect. Caracteristica de corpus deschis, în permanentă îmbogățire, i-ar conferi unicitate între corpusurile de limbă română și i-ar spori astfel valoarea.

Procedura de lucru pe care o avem în vedere nu este spectaculoasă și sperăm să primească idei de îmbunătățire din partea colegilor. Acest corpus nou ar trebui lematizat, astfel încât formele flexionare să fie subsumate unei leme. Procesul nu este lipsit de erori nici pentru limbile cu flexiune mai redusă, precum engleza (Kilgarriiff 2013), dar, chiar și așa, ar reduce munca lexicografului. Următoarea etapă ar constitui-o obținerea unei liste de cuvinte (un index) care să fie comparată automat cu lista de intrări din actuala versiune a DEX, cuvintele deja existente în DEX fiind eliminate⁵. Astfel, redactorii ar rămâne cu o listă de cuvinte neînregistrate în DEX, din care va trebui să elimine erorile grafice (de exemplu, un cuvânt scris greșit ar fi identificat drept cuvânt nou), cuvintele străine (în textele române sunt adeseori citate titluri sau propoziții în engleză) și numele proprii. Din păcate, nu cunoaștem încă o modalitate automată de a le elimina, iar problema „gunoaielor” (eng. *noise*) nu a fost rezolvată nici pentru alte corpusuri ale altor limbi. La final, redactorii ar rămâne cu o listă de cuvinte-candidate pentru includere, pe care trebuie să le evalueze. Pe această listă s-ar mai putea adăuga cuvinte despre care utilizatorii dicționarului au trimis întrebări sau pe care le-au căutat pe site și nu le-au găsit.

În lipsa acestui corpus, DEX ar fi îmbogățit prin metode tradiționale – prin apelul la alte dicționare, precum DCR³ sau mai recentul ICSO, prin excerptarea

⁴ Îi mulțumesc lui Claudiu Teodorescu pentru semnalarea acestei posibilități deja existente pentru rețeaua Twitter.

⁵ Identificarea cuvintelor noi, neînregistrate în dicționare se face și automat în alte limbi (cu instrumente precum Neocrawler, Kerremans *et al.* 2018 sau Neoveille, Cartier 2017), însă limba română nu dispune încă de resursele necesare.

unui număr redus de texte contemporane de către redactori (întrucât posibilitățile de excerptare ale unui om sunt mult reduse față de capacitățile de procesare ale unui corpus precum cel descris mai sus) – și revizia ar avea un caracter mai degrabă empiric și subiectiv. Apelul la un corpus reprezentativ ar aduce DEX³ într-o nouă etapă, conferindu-i un caracter mai bine fundamentat științific.

Odată identificate cuvintele-candidate, acestea ar trebui verificate la fel ca cele deja înregistrate în DEX, atât în corpusul școlar, cât și în corpusul reprezentativ. Prezența oricărui cuvânt în corpusul școlar ar trebui să atragă prezența acestuia în DEX³, indiferent de datele din corpusul de limbă actuală.

Criteriile de incluziune pe care le propunem sunt în linii mari aceleași cu cele menționate în cazul celorlalte dicționare (*vide supra* 3.5). Includerea unui cuvânt în DEX³ ar trebui să fie decisă în funcție de:

(a) numărul atestărilor din corpus: în funcție de dimensiunea corpusului și de dimensiunea dicționarului, ar trebui stabilit un număr minim de atestări care să fie considerat suficient pentru a dovedi un uz stabil. Dacă *Oxford English Dictionary* propune 100 de atestări independente, această cifră trebuie gândită în raport cu resursele (corpus de 2 miliarde de cuvinte plus celelalte resurse, de exemplu Oxford Reading Programme), dar și cu dimensiunea dicționarului. Pentru un dicționar de uz general, numărul atestărilor unice trebuie ponderat cu tipul surselor. Spre exemplu, un termen tehnic poate apărea în literatura de specialitate cu mare frecvență, depășind acel număr minim de atestări, dar poate să nu apară deloc în alte tipuri de texte, situație în care nu poate fi acceptat în DEX³. Numărul minim de atestări este un prim indiciu, dar nu singurul. În lipsa unui corpus reprezentativ de limbă actuală, uzul în limba curentă nu ar putea fi verificat decât prin căutări în internet, setând motorul de căutare numai pentru limba română. Propunerea noastră este un număr minim de 500 de atestări independente, cu aceeași rezervă a necesității de a identifica aceste atestări în texte aparținând mai multor domenii sau stiluri. În lipsa unui alt corpus, strict orientativ ar putea fi interogată și CoRoLa, pentru care propunem un număr minim de 100 de ocurențe independente, verificabile ca aparținând unor domenii diferite.

(b) vârsta cuvântului: propunem un minimum de 5 ani de atestare continuă pentru ca un cuvânt să poată fi acceptat în DEX³. În cazul decuplării site-ului de dicționarul tipărit, vârsta minimă ar putea fi redusă la 2 ani sau chiar mai puțin, în situația unor evenimente sociale acute care determină difuzarea rapidă a unor cuvinte noi, cum este pandemia de SARS-CoV-2. De altfel, dicționarele de uz general din spațiul britanic și francez au inclus imediat în versiunile online, ulterior și în cele tipărite (PR și PL) vocabularul pandemic, considerând că sunt necesare pentru utilizatori, întrucât citesc zilnic știri care conțin aceste cuvinte.

(c) necesitatea la nivel referențial: în cazul unor inovații tehnologice cu difuzare rapidă, denumirea unui obiect foarte folosit ar trebui să se regăsească în DEX³.

(d) relevanța la nivelul sistemului limbii: anumite cuvinte care nu îndeplinesc toate criteriile de mai sus ar putea fi totuși primite dacă sunt necesare pentru explicarea unor derivate sau compuse din familia lor lexicală.

Toate aceste criterii pot ajuta lexicograful să ia decizii informate și bine argumentate, fără a-l obliga. Oportunitatea includerii unui cuvânt în dicționar rămâne în cele din urmă la decizia echipei de redactori, care are însă obligația de a documenta uzul actual al cuvântului. În fine, nu orice cuvânt sau sens care apare cu o oarecare frecvență trebuie inclus în dicționar. Discernământul redactorului înseamnă și recunoașterea unor uzuri nerecomandate și urmărirea lor. De exemplu, *locație* cu sensul „loc” s-a generalizat și a fost acceptat în ediția din 2016 a DEX-ului, dar nu credem că este momentul pentru a include în DEX și alte cazuri de sensuri calchiate după engleză, precum *a salva* cu sensul „a economisi”.

În cele de mai sus, am avut în vedere în primul rând forma tipărită a DEX³, întrucât aici constrângerile de spațiu se manifestă stringent. În perspectivă, credem că este de dorit ca site-ul oficial al DEX să fie decuplat de versiunea pe hârtie. Pe site ar putea fi introduse mai multe cuvinte, chiar dintre cele care nu fac parte din limba curentă, marcate ca atare. Astfel s-ar rezolva și problema arhaismelor și a regionalismelor care se regăsesc în operele literare, dar nu apar și în corpusul școlar reprezentativ, precum și a altor cuvinte din masa vocabularului. În plus, site-ul poate fi actualizat mai des, iar cuvintele și sensurile noi, de interes pentru utilizatori, n-ar trebui să mai aștepte o ediție nouă. Îmbogățirea DEX-ului ar fi un reflex (puțin întârziat) al îmbogățirii corpusului reprezentativ, în cazul în care acest proiect ar fi implementat în cadrul institutului nostru.

5. CONCLUZII

În România, *Dicționarul limbii române* (DEX) e „Dicționarul”. Dintre toate operele lexicografice existente, DEX are cea mai mare autoritate. Elaborarea de către Academia Română îi conferă autoritate: deși DEX este un dicționar în esență descriptiv, este perceput și folosit ca dicționar normativ, iar autoritatea implică și o mare responsabilitate din partea redactorilor. Lista de intrări trebuie alcătuită cu discernământ, ținând cont, pe de o parte, de caracteristicile unui dicționar de uz general și, pe de altă parte, de autoritatea academică a DEX-ului. Odată incluse în DEX, cuvintele, locuțiunile și sensurile sunt validate din perspectiva utilizatorului, care le va considera „corecte”.

Însă alcătuirea listei de intrări ar trebui să evolueze de la metode tradiționale la metode mai moderne, asistate de tehnologie, așa cum s-a petrecut cu multe dicționare din spațiul european. DEX este un dicționar al limbii curente, așadar trebuie să aibă o legătură esențială cu un corpus de limbă curentă, reprezentativ și îmbogățit continuu, corpus a cărui alcătuire care ar trebui să devină proiect fundamental al Institutului de Lingvistică „Iorgu Iordan – Alexandru Rosetti”.

SIGLE

- CB = Cannella, Mario, Roberta Balboni, *Come nasce e come cresce un vocabolario*. Material nedatat, prezent pe site-ul editurii Zanichelli, https://staticmy.zanichelli.it/catalogo/assets/allegati/generici/Come_nasce.pdf, accesat ultima dată la 25.11.2021.
- CADE = I.-A. Candrea, Gh. Adamescu, *Dicționarul enciclopedic ilustrat*. București, Editura Cartea Românească, 1931.
- CED = *Collins English Dictionary*, 13th edition, Glasgow, HarperCollins, 2018, online la <https://www.collinsdictionary.com/us/dictionary/english>
- DA = *Dicționarul limbii române*. Seria veche. Sub conducerea lui Sextil Pușcariu, 1913–1948.
- DCR³ = Florica Dimitrescu, Alexandru Ciolan, Coman Lupu, *Dicționar de cuvinte recente* (DCR³). Ediția a III-a. București, Editura Logos, 2013.
- De Mauro = Tullio De Mauro, *Il dizionario della lingua italiana per il terzo millennio*, Paravia, 2000, online la <https://dizionario.internazionale.it/>
- DEX – S = *Supliment la Dicționarul explicativ al limbii române*. Academia Română, Institutul de Lingvistică „Iorgu Iordan”. București, Editura Academiei, 1988.
- DEX¹ = *Dicționarul explicativ al limbii române*. Academia Română, Institutul de Lingvistică „Iorgu Iordan”. București, Editura Academiei, 1975.
- DEX² = *Dicționarul explicativ al limbii române*. Academia Română, Institutul de Lingvistică „Iorgu Iordan – Al. Rosetti”. Ediția a II-a adăugită. București, Editura Univers Enciclopedic, 1996, retipărit cu adăugiri în 2009, 2012, 2016.
- DEXI = Eugenia Dima (coord.), *Dicționar explicativ ilustrat al limbii române*. Chișinău, Arc & Gunivas, 2007.
- DLE = *Diccionario de la lengua española*, ed. 23, online la <https://dle.rae.es/>.
- DLR = *Dicționarul limbii române*. Serie nouă. București, Editura Academiei, 1965–2010.
- DLRLC = Dimitrie Macrea, Emil Petrovici (coord.), *Dicționarul limbii române literare contemporane*. București, Editura Academiei. Vol. I: A–C, 1955; vol. II: D–L, 1956; vol. III: M–R, 1957; vol. IV: S–Z, 1958.
- DLRM = D. Macrea (coord.), *Dicționarul limbii române moderne*. București, Editura Academiei Române, 1958.
- ICSO = Ana-Maria Barbu, Blanca Croitor, Anabella-Gloria Niculescu-Gorpin, Carmen-Ioana Radu, Monica Vasileanu, *Inventar de cuvinte și sensuri noi atestate în mediul online*. Volumul I. București, Editura Academiei, 2020.
- MW = *Merriam-Webster Dictionary*, online la <https://www.merriam-webster.com/help/faq-words-into-dictionary>
- NDU = Ioan Oprea, Carmen-Gabriela Pamfil, Rodica Radu, Victoria Zăstroiu, *Noul dicționar universal al limbii române*. Chișinău/ București, Litera Internațional, 2006.
- ODE = *Oxford Dictionary of English*, 3rd edition, Oxford, OUP, 2010.
- PL = *Le petit Larousse illustré*, ed. 2022, Editions Larousse. Prefața disponibilă la <https://www.editions-larousse.fr/livre/petit-larousse-illustre-2022-9782035938664>
- PR = *Le petit Robert de la langue française*, ed. 2022, Le Robert. Prefață disponibilă la <https://www.lerobert.com/dictionnaires/francais/langue/dictionnaire-le-petit-robert-de-la-langue-francaise-2022-et-son-dictionnaire-en-ligne-9782321016533.html>
- Prog. = Programa școlară pentru disciplina Limba și literatura română. Clasele a V-a – a VIII-a. București, 2017, online la <http://programe.ise.ro/Portals/1/Curriculum/2017-progr/01-Limba%20si%20literatura%20%20romana%20materna.pdf>
- ȘDU = Lazăr Șăineanu, *Dicționarul universal al limbei române*, Craiova, Institutul de Editură Ralian și Ignat Samitca, 1896.
- Treccani = *Il Nuovo Treccani. Neologismi*, Istituto della Enciclopedia italiana fondata da Giovanni Treccani, Roma, 2018.
- Zingarelli = *Lo Zingarelli* 2021, ed. Mario Cannella, Beata Lazzarini, Andrea Zaninello, Bologna, Zanichelli.

BIBLIOGRAFIE

- Atkins, B. T. S., M. Rundell. *The Oxford Guide to Practical Lexicography*. Oxford; New York: Oxford University Press, 2008.
- Béjoint, H., 2016, „Dictionaries for general users: history and development; current issues”, în Ph. Durkin (ed.), *The Oxford Handbook of Lexicography*. Oxford: Oxford University Press, 7–24.
- Burkhanov, I. Y., 1988, *Lexicography: a Dictionary of Basic Terminology*. Wydawn. Wyższej Szkoły Pedagogicznej w Rzeszowie.
- Cartier, E., 2017, „Neoville, a web platform for neologism tracking”, în A. Peñas, A. Martins (ed.), *Proceedings of the Software Demonstrations of the 15th Conference of the European Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, 95–98.
- Cowie, A. P., „Introduction”, în A. P. Cowie (ed.), 2009, *The Oxford History of English Lexicography*. Oxford: Clarendon Press, I, 1–14.
- Diamond, G., 2016, “Making decisions about inclusion and exclusion”, în Ph. Durkin (ed.), *The Oxford Handbook of Lexicography*, Oxford, Oxford University Press, 532–545.
- Harrison, E., 2014, *Oxford Dictionaries: the Demise of the Printed Edition?*, BBC News, 19.06.2014, <https://www.bbc.com/news/uk-england-oxfordshire-27868836>, accesat ultima dată la 25.11.2021.
- Kerremans, D. et. al., 2018. „Using data-mining to identify and study patterns in lexical innovation on the web: The NeoCrawler”, *Pragmatics & cognition*, 25(1), 174–200.
- Kilgariff, A., 2013, „Using corpora as data sources for dictionaries”, în H. Jackson, *The Bloomsbury Companion to Lexicography*. London; New York: Bloomsbury University Press, 77–96.
- Kosem et al. 2018, „The image of the monolingual dictionary across Europe. Results of the European survey of dictionary use and culture”, *International Journal of Lexicography*, 2018, 1–23.
- Kupietz, M., 2016, „Constructing a corpus”, în Philip Durkin (ed.), *The Oxford Handbook of Lexicography*. Oxford: Oxford University Press, 62–75
- Marello, C. (2020). „New words and new forms of linguistic purism in the 21st century: the Italian debate”, *International Journal of Lexicography*, 33(2), 168–186.
- Maxwell, Kerry, 2006. New word of the month. *MED Magazine*, issue 38, online la <http://macmillandictionaries.com/MED-Magazine/May2006/38-New-Word.htm>
- van Sterkenburg, P., 1996, „‘The’ dictionary: Definition and history”, în P. van Sterkenburg, *A Practical Guide to Lexicography*. Amsterdam; Philadelphia: John Benjamins, 3–17.
- Swanepoel, P., 1996, „Dictionary typologies: A pragmatic approach”, în P. van Sterkenburg, *A Practical Guide to Lexicography*. Amsterdam; Philadelphia: John Benjamins, 44–69.

**CRITERIA FOR THE HEADWORD LIST OF A DICTIONARY FOR GENERAL USERS:
SUGGESTIONS FOR DEX³**

Abstract

DEX, the dictionary for general users elaborated by the Romanian Academy, is supposed to undergo a major revision in the forthcoming years. The previous editions contained no information about the criteria for building the headword list. Our article seeks to formulate such criteria, by looking at the characteristics of general-purpose dictionaries, on the one hand, and by analysing several such dictionaries published in Europe, on the other hand. In order to assess the true diffusion of words in current Romanian, the lexicographers need two important tools: (a) a corpus made of all textbooks used in secondary and high schools; (b) a new representative corpus of current Romanian texts, that would be augmented every year. This corpus should become one of our institute's fundamental projects.